

Die folgende Betrachtung einiger Implikationen der Debatte über künstliche Intelligenz wirft einen „fremden“ Blick auf das Expertenuniversum der KI-Forschung. Voraussetzung dafür ist zunächst ein Blick zurück.

Das prometheische Bedürfnis, Menschen nicht bloß zu reproduzieren, sondern sie sogar, wie die Götter, produzieren zu können, findet in der Neuzeit einen säkularisierten Ausdruck: als Bedürfnis, intelligente Maschinen zu erzeugen. Doch bereits in der Antike ist der utopische Wunsch formuliert worden, Maschinen nicht nur als Werkzeuge, sondern als selbständig arbeitende und denkende Produktionseinheiten einsetzen zu können. Für Aristoteles stand es fest, daß man nicht geistig und künstlerisch arbeiten könne, wenn es keine Sklavenarbeit gäbe, die von der lebensnotwendigen körperlichen Arbeit befreit. Darum log er die Sklaverei zu einer Naturtatsache um. Der Sklave sei nichts anderes als ein „beseeltes Werkzeug“. ¹ Aber Aristoteles faßte eine Möglichkeit in den Blick, wie wir Muße auch ohne Sklavenarbeit haben könnten: Nämlich dann, wenn Maschinen von allein die notwendige physische Arbeit verrichten würden. Das lag für ihn jedoch im Bereich der Mythologie. ² Er wollte demonstrieren, wie absurd es wäre, sich eine Gesellschaft ohne Sklaven vorzustellen. Aber damit dachte er zugleich den Gedanken, den er abwehren wollte. Wenn die Werkzeuge von selber arbeiten könnten, würde die Freiheit ihrer Hersteller vergrößert. Und wenn sie auch noch die Befehle „schon vorausahnen“ ³ könnten, wie Aristoteles schreibt, dann würden sie uns noch mehr entlasten.

Die Neuzeit hat sich angeschickt, zu verwirklichen, was für Aristoteles utopisch war. Aber in dem Maße, wie Maschinen als Werkzeuge industrieller Produktion der Lebensbedingungen der Menschen in den Vordergrund traten, mußte man feststellen, daß unter den Prämissen des neuzeitlichen naturwissenschaftlichen Weltbilds kaum noch zwischen Mensch und Maschine zu unterscheiden ist. Das aber läuft der Intention zuwider, die menschliche Freiheit zu vergrößern. Wenn es sich gezeigt hat, daß man nicht vernünftig beweisen kann, daß der Mensch als Geistwesen das Ebenbild Gottes ist, dann bekommt das klassische Leib-Seele-Problem in der Philosophie eine neue, geradezu beängstigende Dimension. Die Physis des Menschen läßt sich nun ohne weiteres als Maschine beschreiben. Für Descartes war es noch lediglich eine methodische Fiktion, die Menschen als Maschinen anzusehen, die sich von selbst bewegen. ⁴ Im Konzept des Maschinenmenschen wurde dann aber die Analogie von Natur und Artefakt, von Organismus und Technik auch noch auf den Menschen selbst ausgedehnt. Es gibt nicht eine materielle, körperliche Substanz und eine spirituelle, denkende, sondern nur noch eine einzige. Für den mechanischen Materialismus des atheistischen Aufklärers La

Mettrie ging der Geist des Menschen daher restlos aus der Materie seines Körpers hervor. ⁵

Wenn Geist auf einen mechanischen Fortsatz materieller Prozesse reduziert ist, läßt sich das Problem angehen, wie man mechanische Artefakte zum Denken bringen könnte. Denken soll dabei zu einer neuen qualitativen Dimension vorstoßen können, denn durch Mechanisierung wäre es möglich, Fehler zu eliminieren und das Verhältnis von Aufwand und Ertrag zu optimieren. Das läßt sich freilich nur dann in den Blick nehmen, wenn man unter Denken instrumentelles Denken versteht. Instrumentelle Rationalität ist, mit den Worten von Charles Taylor, „die Art von Rationalität, auf die wir uns stützen, wenn wir die ökonomischste Anwendung der Mittel zu einem gegebenen Zweck berechnen.“ ⁶

Die Schranke künstlicher Intelligenz wäre eine Denkmaschine, die sich von den Bedingungen der Möglichkeit ihrer Tätigkeit unabhängig machen würde, und zwar in doppelter Hinsicht. Sie müßte imstande sein, selbständig und kreativ gedankenähnliche Prozesse in Gang zu setzen. Und sie müßte womöglich ihren eigenen Charakter als Artefakt erkennen und sich aufgrund dessen selbst programmieren können. Diese Grenze ist noch nicht erreicht. Aber wenn eine Grenze erst einmal gesetzt ist, dann wissen wir bereits, was auf ihrer anderen Seite liegt und haben dieses Jenseits damit auch schon ein Stück weit im Griff.

Der Traum von selbständig *arbeitenden* Maschinen ist erfüllt – und dabei zugleich zum Alptraum geworden. Der Traum von selbständig *denkenden* Maschinen rückt seiner Erfüllung näher – und zwar in dem Maße, wie sich das Denken der Menschen den Rechenoperationen der Computer anähnt. Joseph Weizenbaums Dialogprogramm „Eliza“ setzte die stereotypen Phrasen der Gesprächspsychotherapie von Carl Rogers raffiniert ein: als Reiz-Reaktionsschemata menschlicher Kommunikation. Das Programm sollte den Mythos zerstören, daß wir eine selbstreferentielle, autonome künstliche Intelligenz herstellen können. „Eliza“ führte vor, wie leicht der Schein reflektierter Kommunikation erzeugt werden kann, wo doch in Wirklichkeit nichts anderes passiert, als daß Antwort- mit Fragestrukturen verknüpft werden, anhand formal definierter Merkmale von Satzkonstruktionen. Der Entlarvungsversuch scheiterte. Erstens, weil sich Therapie-Erfolge bei Probanden einstellten, die nicht wußten, daß sie mit einer Rechenmaschine kommunizierten. Und zweitens vor allem deshalb, weil der Unterschied zwischen einem behavioristischen Begriff von Intelligenz und einem, der philosophische Implikationen hat, aus dem Blick geraten ist.

In der Expertenwelt der Künstlichen-Intelligenz-Forschung geht es um die Frage, ob Computer denken können oder nicht. Um sie beantworten zu können, wird definiert, was *Denken* heißt. Allen Newell und Herbert Simon verstehen unter Denken

die „Verarbeitung von Informationen“. Diese wird als „Rechenvorgang“ aufgefaßt, bei dem Zeichen, „Symbole“, in bestimmte Beziehungen zueinander gebracht werden. Eine Verbindung zur außersymbolischen Wirklichkeit könne vorausgesetzt werden, weil Symbole „Bedeutungen“ haben,⁷ d.h., sich auf etwas beziehen, was sie nicht selbst sind. So verstanden, ist Denken ein rein formaler Vorgang der regelgeleiteten Verknüpfung und Manipulation von Zeichen; ein Instrument, dessen sich intelligente Wesen bedienen. Aber was wird unter *Intelligenz* verstanden? Als ihr wesentliches Merkmal definiert Daniel C. Dennett „die Fähigkeit, zu lösende Probleme zu erkennen und klar über sie nachzudenken.“⁸

Viele Forscher glauben, es sei möglich, Rechenmaschinen zu konstruieren, die in der Lage wären, „zu lösende Probleme zu erkennen und klar über sie nachzudenken“, indem sie „Symbole in einem Rechenvorgang verknüpfen“. Sie vertreten übereinstimmend folgende Ansicht: Wenn es einem Rechner gelingen würde, einen geschickt fragenden menschlichen Gesprächspartner „regelmäßig oder oft“⁹ darüber zu täuschen, daß er sich nicht mit einem anderen Menschen, sondern mit einem Computer unterhält, dann muß von diesem Computer gesagt werden, daß er intelligent sei und denken könne. Bislang hat noch kein Computer diesen „Turing-Test“ unter strengen Bedingungen bestanden. Interessanter als die Diskussion, ob dies eines Tages machbar wäre, ist aber die Betrachtung des Intelligenzbegriffs, der hier zugrunde gelegt wird. Der Begriff von Intelligenz als instrumentelles Vermögen, denkend Probleme zu lösen und sich an die Umwelt anpassen zu können, stammt aus dem Behaviorismus. Für die Philosophie, die sich, wie ich mit den Worten von John Searle sagen möchte, von der „Tradition einer unverfornen behavioristischen und operationalistischen Haltung“¹⁰ abgrenzen sollte, ist Intelligenz aber mehr als das. Sie ist ein Bestandteil menschlicher Rationalität, der untrennbar verbunden ist mit bestimmten anderen Bestandteilen: Spontaneität, Selbstbezüglichkeit, Kritikfähigkeit und Erfahrung.

Spontaneität ist, im Sinne Kants, das Vermögen des Verstandes, »Vorstellungen« eines Gegenstandes „selbst hervorzubringen“¹¹, also die Fähigkeit zur Begriffsbildung, letztlich: zur Erkenntnis. *Selbstbezüglichkeit* bedeutet Selbstreflexion, aber auch *Selbstkritik* des Denkens. „Selbstkritik der Vernunft“, so Herbert Schnädelbach, ist das „Gegenmittel gegen ihr eigenes Blind- und Dummwerden, gegen ihre Selbstzerstörung und Selbstaufhebung“¹² – also etwas, ohne das es überhaupt keine Fähigkeit zur Kritik gäbe. Auch der Begriff der *Erfahrung* zeigt, mit welchen Ausgrenzungen das Konzept der künstlichen Intelligenz arbeiten muß. Denken als geistige Erfahrung steht in einem lebendigen Austauschprozeß mit seinen Gegenständen. Diese Dimension geht

verloren, wenn Bewußtsein, Intelligenz und Rationalität nur *formal* definiert werden.

Intelligenz ist also lediglich *ein Moment* dessen, was in der Philosophie als Geist oder Vernunft bezeichnet wurde. Der gegenwärtige Begriff künstlicher Intelligenz zeigt, daß sich die Auffassung vom menschlichen Geist, wie sie für die neuzeitliche Kultur prägend gewesen ist, gewandelt hat. Heute wird nicht mehr der menschliche Geist als Vorbild des maschinellen Verknüpfens von Daten gedacht, das dann folgerichtig als technologische Quasi-Geistigkeit aufzufassen ist. Umgekehrt: Die formale, quantitative Leistung von Maschinen ist zum neuen Grundmuster des Geistes geworden. Mit den Worten von Sybille Krämer: „Seit der Mitte unseres Jahrhunderts [...] werden Untersuchungen über den Geist formuliert in den Termini von Informationsverarbeitung, deren Vorbild zumeist gewonnen wird an der Datenverarbeitung durch den Computer. Der Aufstieg des Informationsverarbeitungsmodells in nahezu allen Wissenschaften, die das Denken untersuchen, signalisiert, daß der anthropozentrische Geistbegriff der Neuzeit in Auflösung begriffen ist und daß er abgelöst wird durch einen [...] maschinenzentrierten Begriff vom Geist.“¹³

Doch dieser Begriff des Geistes steht im Widerspruch zu sich selber. Searle hat mit seinem bekannten Bild des „Chinese Room“¹⁴ das Defizit nachgewiesen: Die Vorstellungen von *Bedeutung* und *Verstehen*, die diesem Begriff des Geistes zugrundeliegen, werden rein formal definiert. Ein Computer kann mit gegebenen Symbolen nach gegebenen Regeln operieren, aber die Symbole haben *für ihn* keine *Bedeutung*, sie verweisen auf nichts, das außerhalb ihrer selbst liegt, und daher versteht der Computer auch nicht, was er tut. Mit diesem Argument ist der Fehler in der formalen Bestimmung des Denkens bei Newell und Simon offengelegt: Die Verbindung zwischen Symbol und Wirklichkeit ist für den Computer eben nicht gegeben. Er operiert wie ein Mensch, der in einem verschlossenen Zimmer sitzt und Karteikarten mit chinesischen Schriftzeichen einander zuordnet, nach Regeln, die ihm vorgegeben werden, ohne daß er die chinesische Sprache kennt. Wenn die Regeln sinnvoll und in einer Sprache verfaßt sind, die der Mensch versteht, dann wird es ihm gelingen, die Schriftzeichen z. B. zu einer Geschichte zu verknüpfen, die er dann als Output aus dem Zimmer reichen kann. Aber er wird nicht *verstanden* haben, was er so zusammengefügt hat. Er hat „abstrakte Symbole nach formalen Regeln“¹⁵ manipuliert. Er könnte aber mehr, als er hier getan hat: Ein Mensch kann Zeichen, die für ihn bedeutungsvoll sind, zu neuen sinnvollen Verknüpfungen zusammenfügen, weil er über *Intentionalität* verfügt. Darin besteht nach Searle die einzigartige Qualität des menschlichen Bewußtseins. Unsere organische Substanz, die sich biologisch, physikalisch und chemisch be-

schreiben läßt, ist gleichzeitig „unter bestimmten Bedingungen imstande [...], Wahrnehmen, Handeln, Verstehen, Lernen und andere intentionale Zustände hervorzurufen.“¹⁶

Die Prozesse in unserem Organismus lassen sich naturwissenschaftlich *beschreiben*, denn sie sind materielle Vorgänge. Doch der Übergang von Materie in geistige Prozesse läßt sich nicht aus den materiellen Prozessen heraus kausal *erklären*. Denkvorgänge lassen sich zwar als formale Rechenvorgänge simulieren, aber die Simulation ist *nicht identisch* mit dem Simulierten. Und von Bewußtsein, das Intentionalität aufweist, kann nur gesprochen werden, sofern es auch mit allen körperlich-seelischen Eigenschaften verbunden wäre, die der Mensch aufweist und die ihn in die Lage versetzen, wahrnehmen, handeln, verstehen und lernen zu können. Der Begriff der *Intelligenz* muß also im Kontext des Begriffs des *Bewußtseins* (oder des *Geistes*) und des *Leibes* diskutiert werden. Dann zeigen sich die Defizite, die aus seiner formalistischen Verkürzung folgen.

Der „maschinenzentrierte Begriff vom Geist“ steht allerdings für eine Entwicklung, die noch weiter geht. Immer selbstbewußter werdend, hatte unsere geistige Tradition den Geist einst aus der über- oder außermenschlichen, der kosmologischen oder göttlichen Sphäre herausgenommen; sie hatte ihn den Menschen als ihr spezifisches Gattungsvermögen zugeeignet. Heute wird der Geist den Menschen von der Wissenschaft wieder genommen und in eine Ordnung verlegt, die sich ihrem Zugriff immer mehr entzieht. Das „Projekt der Moderne“ war durch die Überlegung gekennzeichnet, daß Produkte menschlicher Tätigkeit, die den Menschen aber nicht als ihr eigenes Werk erscheinen, sondern als Naturverhältnisse, die er hinzunehmen habe, ihrer fetischhaften Erstarrung beraubt werden sollten. Sie sollten als veränderbar erwiesen werden. Entzauberung, Befreiung, Autonomie, ungegängerter Gebrauch der Rationalität – die Freiheit der Menschen war das Ziel, dem Naturwissenschaft zu dienen hatte. Die *Subjektivität* ist, mit den Worten von Jürgen Habermas, das „Prinzip der Moderne“, verstanden „als uneingelöstes Versprechen“. Naturwissenschaft und Technologie sind funktional auf dieses Prinzip bezogen. Ihr philosophisch plausibler Status wäre der von Hilfsmitteln einer „selbstbewußten Praxis“, „in der sich die solidarische Selbstbestimmung aller mit der authentischen Selbstverwirklichung eines jeden einzelnen verbinden können.“¹⁷ Diese normativen Grundlagen der Moderne sind in der Erforschung und dem technologischen Einsatz künstlicher Intelligenz nicht mehr zu entdecken. Der aufklärerische Nachweis, daß es keine fixen Naturverhältnisse gibt, daß alles veränderbar ist und von den Menschen selbst hergestellt werden kann, wird zwar geleistet, aber dabei zugleich in seiner Intention umgekehrt.

Naturbeherrschung, die immer noch gesellschaftlich unbeherrscht ist, und Wissenschaft, der philosophische Begriffe und kritische Selbstreflexion immer mehr abhanden zu kommen scheinen, tendieren zum programmatischen Verzicht auf Autonomie. Der Orientierungsbedarf wird statt dessen durch die freiwillige Abdankung des Geistes befriedigt. „Erst Gott, dann der Mensch und schließlich die Maschine werden zu den sich ablösenden Leitbildern, an denen die Antike, die Neuzeit und schließlich die Moderne ihren Entwurf vom Geist jeweils ausrichten.“¹⁸ Das Bedürfnis nach Autonomie wird heute nicht mehr als praktisches Interesse begriffen, das gesellschaftlich zu realisieren wäre. Es erscheint als hybrider Urzeugungswahn, der im Denken und in der Praxis in verschiedenen Gestalten aufgetreten ist: als glücklicherweise überwundene Ideen wie Gottesebenbildlichkeit oder Anthropozentrik. Von dem Projekt, in freiem Triumph über die vorgefundenen Bedingungen naturhafter Reproduktion Menschenähnliches zu erzeugen, bleibt unter den Bedingungen einer antagonistischen Konkurrenzgesellschaft nichts mehr übrig. Im Zeichen der Resignation von der regulativen Idee der Moderne erscheint es schließlich, als könne gesellschaftliche Autonomie gar nichts anderes sein als die in die Totale projizierte Hybris, als eine Veranstaltung, die theoretisch wahnhaft ist und praktisch verfehlt.

Die künstlich-intelligenten Maschinen haben das Gewünschte bereits verwirklicht. Sie sind Vervollkommnungen menschlicher Intelligenz und sind gleichzeitig moralisch stabil. Sie können nahezu alles ersinnen, was Menschenhirne ersinnen können, und sie werden dabei nicht von moralischen Schwankungen, Verlockungen oder Zweifeln gefährdet. Das funktioniert, weil sie die Potenzierung eines verkürzten Begriffs von Intelligenz und daher auch moralisch indifferent sind. Sie sind fähig, „Symbole in einem Rechenvorgang zu verknüpfen“, um auf diese Weise „zu lösende Probleme zu erkennen und klar über sie nachzudenken“. Insofern sind sie eben doch stets die Werkzeuge ihrer Produzenten. Aber diese haben sich ihren Produkten angeglichen. Wir verlangen anscheinend gar nicht mehr von uns, als *solche* „Probleme zu erkennen und klar über sie nachzudenken“, die sich auf Symbole reduzieren lassen, welche »in einem Rechenvorgang verknüpft« werden können. Andere Probleme geraten aus dem Horizont. Der „Sklavenaufstand der Technik“¹⁹ war erfolgreich. Der Traum von der künstlichen Intelligenz realisiert sich – aber nur in Gestalt des Verzichts auf Selbstreflexion. Der Maschinenmensch ist nicht, wie La Mettrie dachte, etwas Naturhaftes. Er kommt erst als spätes gesellschaftliches Resultat zu sich selbst, indem er sich zugleich anschickt, sich als Menschen zu verleugnen.

Hinter den erreichten Stand des technologischen

Wissens und Könnens könnten wir nur um den Preis der kulturellen Regression zurückgehen. Über die Zweckbestimmung der technologischen Zivilisation als solche aber können die Werkzeuge selbst nicht befinden. Auch die technologische Rationalität als ganze kann das nicht – ganz gleich, ob sie „künstlich“ oder „natürlich“ ist. Und vor allem kann sie die Reflexion auf Zwecke nicht autonom hervorbringen und begründen.

Die christlich-abendländische Idee der Gotesebenenbildlichkeit und ihre säkularisierten, aufgeklärten Transformationen sind in der Moderne neu formuliert worden: als praktischer Imperativ, die Welt vernünftig einzurichten, sie zum Lebensraum freier und solidarischer Menschen zu machen. An die Stelle dieses Imperativs ist heute ein anderer getreten. Es ist der praktizierte Imperativ der Maschinenebenenbildlichkeit. Damit wird sich aber auf lange Sicht vielleicht sogar noch die reduzierte Bestimmung von Intelligenz selbst durchstreichen. Denn wenn über die Zwecke, denen instrumentelle Vernunft sich unterordnet, überhaupt nicht mehr rational befunden wird, dann können sie sich selbständigen. Dann können Sach- und Verwertungszwänge zu quasi-naturwüchsigen Kräften werden, die unsere Lebensinteressen nachhaltig gefährden. Davon bekommen wir eine Ahnung, wenn wir uns die waffentechnologische und ökonomische Bedeutung der künstlichen Intelligenz vergegenwärtigen. „Das Aufkommen intelligenter Maschinen“, frohlocken ihre Apologeten, „verändert [...] die weltweiten Handelsbeziehungen.“ Auch die Profiteure der Tötungstechnologie sind dabei auf sie angewiesen. „Militärische Opera-

tionen wie die Zerstörung russischer SAM-Stellungen in Syrien durch Israel oder der Einsatz 'intelligenter' Raketenköpfe im Falklandkrieg haben gezeigt, daß auch im Bereich der Verteidigung die Bedeutung künstlicher Intelligenz wächst. Viele Militärbeobachter sagen voraus, daß in den neunziger Jahren die KI-Technologie von größerer strategischer Bedeutung sein wird als Truppenstärke, Geographie oder natürliche Ressourcen“²⁰, kommentiert enthusiastisch eine US-amerikanische Publikation, die kurz vor dem Golfkrieg erschien.

Intelligenz, verstanden als Anpassung an die gegebene Situation, sollte Selbsterhaltung gewährleisten. Sie kann, so verstanden, aber zur Triebfeder des Selbstverlusts werden. Insofern mag es den Anschein haben, als würde künstliche Intelligenz immer mehr als künstlicher *Intelligenzverzicht* der Menschen realisiert. Das Projekt der Selbstbesinnung der Rationalität auf ihren humanen Zweck verträgt sich mit künstlicher und instrumentell verkürzter Intelligenz; denn es widerspricht ihr keineswegs. Aber es bedarf der intentionalen, selbstreflektierten Intelligenz, also der Vernunft, wie sie nur in menschlichen, leibhaften Subjekten vorkommt, um diese Zielbestimmung zu formulieren und vielleicht zu verwirklichen. Ohne Selbstbesinnung keine Selbstbestimmung. Andernfalls werden wir von den technischen Mitteln versklavt, die doch bei der Abschaffung der Sklaverei helfen sollten.

Verfasser:

*Dr. phil. Gerhard Schweppenhäuser
Bauhaus-Universität Weimar*

Anmerkungen

- 1 Aristoteles, Politik, 1254 a 25.
- 2 Vgl. Aristoteles, a. a. O., 1254 a 30 ff.
- 3 Vgl. ebd.
- 4 Vgl. René Descartes, Meditationen, Hamburg 1972, S. 25.
- 5 Vgl. Julien Offray de La Mettrie, Der Mensch eine Maschine, Leipzig 1984, S. 97 ff.
- 6 Charles Taylor, Das Unbehagen an der Moderne, Frankfurt a. M. 1995, S. 11.
- 7 Vgl. Mitchell Waldrop, Können Computer denken, in: R. Kurzweil, Das Zeitalter der Künstlichen Intelligenz, München/ Wien 1993, S. 63.
- 8 Daniel C. Dennett, Können Maschinen denken?, in: Kurzweil, a. a. O., S. 50.
- 9 Dennett, a. a. O., S. 49.
- 10 John Searle, Geist, Gehirn, Programm, in: Künstliche Intelligenz. Philosophische Probleme, hg. v. W. Ch. Zimmerli u. S. Wolf, Stuttgart 1994, S. 261 f.
- 11 Immanuel Kant, Kritik der reinen Vernunft, B 75.
- 12 Herbert Schnädelbach, Vernunft, in: Philosophie, hg. v. H. Schnädelbach u. E. Martens, Reinbek bei Hamburg 1985, S. 92.
- 13 Sybille Krämer, Künstliche Intelligenz – Der Aufstieg des Computers zum Modellbaukasten des Geistes, in: Information Philosophie Mai 1994, S. 16.
- 14 Vgl. Searle, a. a. O., S. 232 ff.
- 15 Waldrop, a. a. O., S. 54.

- 16 Searle, a. a. O., S. 255.
- 17 Jürgen Habermas, Der philosophische Diskurs der Moderne, Frankfurt a. M. 1986, S. 391.
- 18 Krämer, a. a. O.
- 19 Walter Benjamin, Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit, in: ders., Gesammelte Schriften Bd. I.2, Frankfurt a. M. 1980, S. 469.
- 20 Raymond Kurzweil, Das Zeitalter der Künstlichen Intelligenz, a. a. O., S. 8 u. 9.